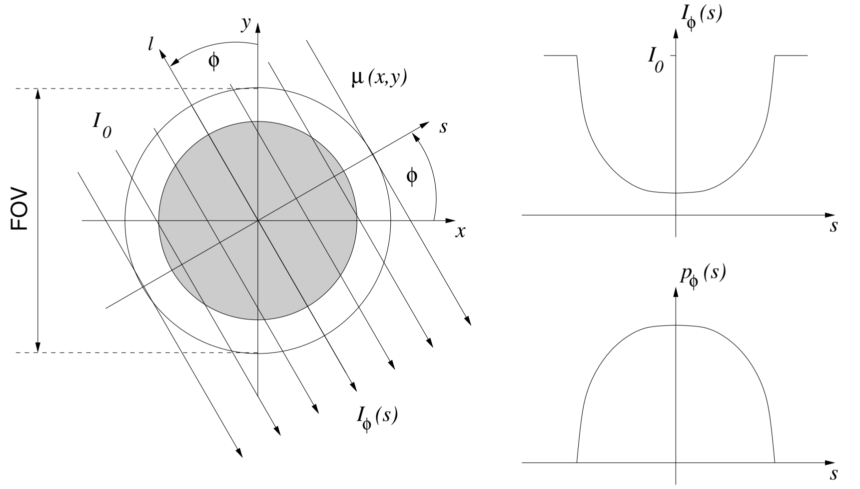
Engineering Tripos Part IIA THIRD YEAR

Module 3G4: Medical Imaging & 3D Computer Graphics

Solutions to 2025 Tripos Paper

1. **CT sampling and aliasing**

(a) A projection at angle ϕ is the set of all line integrals through the function perpendicular to a line which makes an angle ϕ with the x axis.



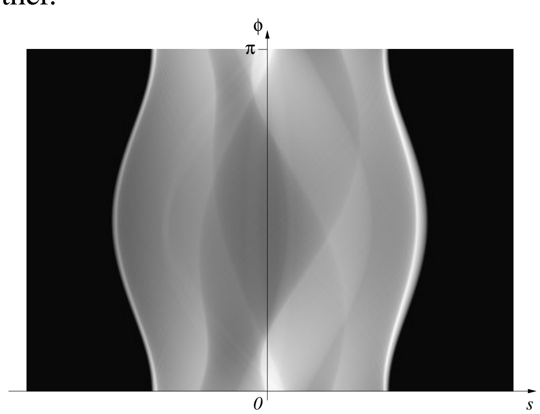
If $\mu(x, y)$ is a function defined on the (x, y) plane, then

$$p_\phi(s) = \int_{-\infty}^{+\infty} \mu(s\cos\phi - l\sin\phi, s\sin\phi + l\cos\phi)\, dl$$

is the projection of μ at an angle ϕ.

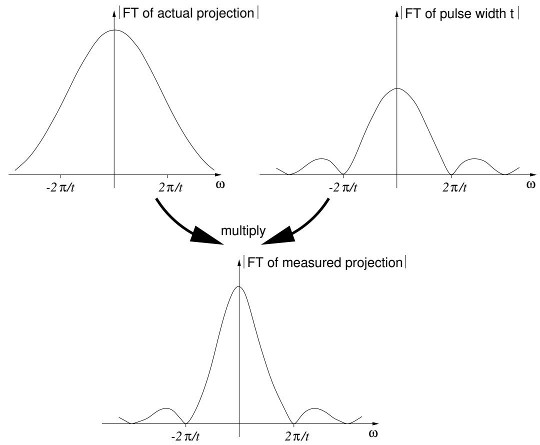
A projection may be displayed with amplitude encoded as brightness.

A sinogram is the set of all the projections of a single function, at different angles, stacked up one on top of another.



[25%]

(b) See the *Information Data Book* for the Fourier transform of a pulse. The first zero is at $2\pi/t$.



The convolved projection is then sampled at a rate Δs, resulting in the spectrum repeating every $2\pi/\Delta s$. To limit aliasing, we should separate the spectra as far as possible, at least until the first zero-crossings coincide. Thus $2\pi/\Delta s \geq 4\pi/t$, giving $\Delta s \leq t/2$. [25%]

(c) The steps shown in Fig. 1 arise from the projection theorem, which states that the 1D Fourier transform ($\mathcal{F}_{1(\omega)}[\,]$) of the projection data $p_\phi(s)$ at a given projection angle ϕ is the same as the radial data passing through the origin at a given angle ϕ in the 2D Fourier transform ($\mathcal{F}[\,]$) of the attenuation data $\mu(x, y)$.

$$\mathcal{F}_{1(\omega)}\left[p_\phi\right] = \mathcal{F}\left[\mu\right]$$

Using the projection theorem, the attenuation data can be reconstructed as follows.

- Take n projections $p_\phi(s)$ of the object. In practice, if there are m X-ray detectors, we will have only m samples of $p_\phi(s)$.
- Find 1D Fourier transforms of the projections. In practice, this will be a discrete Fourier transform resulting in m data points in the frequency domain.
- Use the set of 1D Fourier transforms to tile the spatial frequency plane. Each transform contributes a radial strip (ω, ϕ), for a particular angle ϕ, passing through the origin.
- Resample this data to produce a regular sampling of the spatial frequency plane in Cartesian coordinates (ω_x, ω_y). This involves interpolation to fill the gaps between the radial strips.
- Take the 2D inverse Fourier transform.

The first three steps are illustrated in Fig. 1, with the final two steps completing the process. [25%]

(d) Let the number of projections be n, the number of X-ray detectors be m and the highest spatial frequency measured for each projection (i.e. the length of the "spokes" in Fig. 1) be ω_{max}. Then:

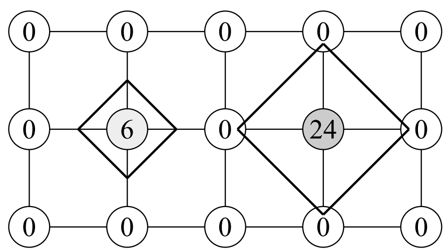
$$\Delta_1 \approx \frac{2\omega_{\max}}{m} \approx \Delta_2 = \frac{\pi\omega_{\max}}{n} \Rightarrow n \approx \frac{\pi m}{2}$$

So the number of projections should be approximately the same as the number of X-ray detectors. [25%]

Assessor's remarks: This question tested the candidates' understanding of computed tomography and direct Fourier reconstruction. The book work in parts (a) and (c) was well understood by most, though a few candidates thought that the Ram-Lak filter played a part in direct Fourier reconstruction. Part (b) was poorly answered, which was surprising seeing as it was also asked on the examples paper. The only truly unseen part of the question was (d), in which candidates were asked to consider how many projections are required. Although there were several perfect answers, most candidates struggled to get started with this simple, two-line analysis.

2. **Marching squares and interpolation**

(a) Marching squares uses linear interpolation of data values to work out where the object contour crosses each line on the grid, then joins up these points within each square. The resulting contour is as below.



(b) (i) B-spline approximation is separable so 2D approximation is most easily achieved by approximating first in one dimension, then repeating in the other dimension. B-spline approximation in one dimension is achieved using:

$$v = \left[\begin{array}{cccc} t^3 & t^2 & t & 1 \end{array}\right] \frac{1}{6} \left[\begin{array}{rrrr} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{array}\right] \left[\begin{array}{c} v_{-1} \\ v_0 \\ v_1 \\ v_2 \end{array}\right]$$

where v is the approximated value and $v_{-1} \ldots v_2$ are the input values. As t varies from 0 to 1, this will generate output values for locations between v_0 and v_1. But here we are only interested in the output actually at the sample location v_0, in which case $t = 0$. Hence:

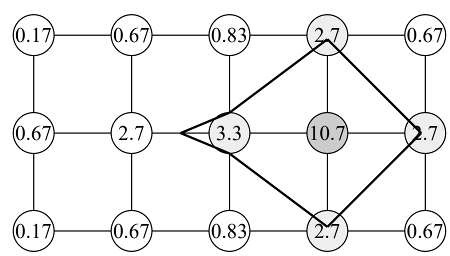
$$v = \frac{1}{6} \left[\begin{array}{ccc} 1 & 4 & 1 \end{array}\right] \left[\begin{array}{c} v_{-1} \\ v_0 \\ v_1 \end{array}\right]$$

If we apply this both horizontally and vertically, we have:

$$\mathbf{W} = \frac{1}{6}\begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} \frac{1}{6}\begin{bmatrix} 1 & 4 & 1 \end{bmatrix} = \frac{1}{36}\begin{bmatrix} 1 & 4 & 1 \\ 4 & 16 & 4 \\ 1 & 4 & 1 \end{bmatrix}$$

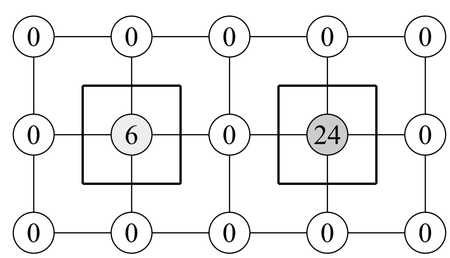
as required.

(ii) We can apply $\mathbf{W}$ to the given data by centring it on each sample (and noting that all other samples are zero). Doing this and then applying marching squares at three gives:



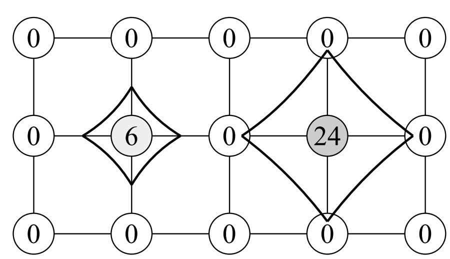
Note that the left-hand non-zero value is now below this threshold.

(c) (i) Nearest neighbour interpolation followed by marching squares gives:



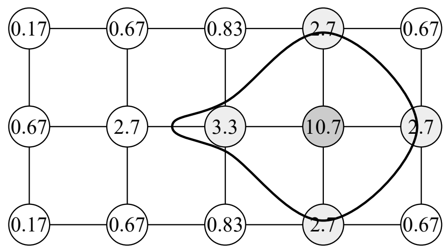
Note that the contours are both square and the same size, since both values 6 and 24 are copied the same distance across the higher resolution grid: the value of 3 only affects the linear interpolation for marching squares over the very small squares in this new grid.

(ii) Bi-linear interpolation followed by marching squares gives:



This is almost the same as the original contour, (since marching cubes was already using linear interpolation). However bi-linear interpolation generates a slightly curved contour: marching cubes only used linear (not bi-linear) interpolation over the edges of the original grid.

(iii) B-spline interpolation followed by marching squares gives (approximately):



This is just a smoothed version of the contour shown in (b)(ii).

(d) Whilst the original object shape is of course not known, the initial sketch in (a) seems likely to be the most accurate, since it is split in two regions and each region has an appropriate size and is not concave.

Smoothing the data first as in (b) joins the two regions, makes the size of the right-hand one larger, and completely misses the left-hand sample which is now outside the contour.

Nearest neighbour interpolation in (c)(i) creates two regions which are the same size, though they are both a sensible shape. Given the large difference in data values, this does not seem likely.

Bi-linear interpolation in (c)(ii) is very similar to (a) and hence nearly as good, but the concave shape does not seem likely.

B-spline interpolation in (c)(iii) creates a smooth contour which does seem more likely, but has joined the two regions together as in (b) which doesn't.

Assessor's remarks: This was a question mainly on interpolation and approximation of scalar data, as well as the use of Marching Squares. There were many good answers, including some very accurate sketches of the differences between the techniques. Most students made good attempts at (a) and (b)(ii). The proof in (b)(i) was well answered by some but there were also quite a few answers which missed the key point that the B-spline only needed to be evaluated at $t = 0$. A surprisingly large number of students did not notice that (c)(iii) was just a smoothed version of the answer to (b)(ii) and hence lost marks here even though their previous answer was correct. It was pleasing to see how many students looked carefully at their sketches and hence came up with good answers to (d). Less pleasing were the generic answers about B-splines being smoother which did not reference the actual results.

3. **Laser scanning**

(a) (i) If we imagine that the laser reflection is located at a distance x from the centreline of the image array, and this makes an angle α with the focal point and centre of the image, then:

$$
\begin{aligned}
\tan \alpha &= \frac{x}{f} \\
\tan(\theta + \alpha) &= \frac{L}{Z} \\
Z &= \frac{L}{\tan(\theta + \alpha)} \\
&= \frac{L(1 - \tan\theta \tan\alpha)}{\tan\theta + \tan\alpha} \\
&= \frac{L\left(1 - \tan\theta \frac{x}{f}\right)}{\tan\theta + \frac{x}{f}} \\
&= \frac{L(f - x\tan\theta)}{f\tan\theta + x}
\end{aligned}
$$

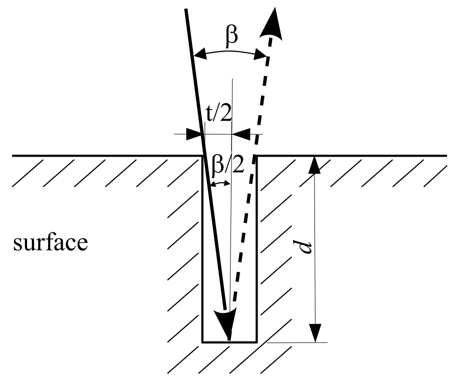
Evaluating this with the minimum and maximum values of x of $-\frac{w}{2}$ and $\frac{w}{2}$ and noting that $\tan\theta = \frac{1}{\sqrt{3}}$:

$$
\begin{aligned}
Z_{\text{min}} &= \frac{30\left(3\sqrt{3} - 1\right)}{3 + \sqrt{3}} \\
&= 26.6\,\text{cm} \\
Z_{\text{max}} &= \frac{30\left(3\sqrt{3} + 1\right)}{3 - \sqrt{3}} \\
&= 146.6\,\text{cm}
\end{aligned}
$$

(ii) For the maximum depth, we need the minimum angle β between the laser light and reflection, which will be at the maximum distance $Z = 146.6\,\text{cm}$. At this distance, the angle between the laser and reflection is given by:

$$
\begin{aligned}
\tan\beta &= \frac{L}{Z} \\
\beta &= 11.57^\circ
\end{aligned}
$$

For maximum depth into the hole, the orientation will be as below.



Hence:

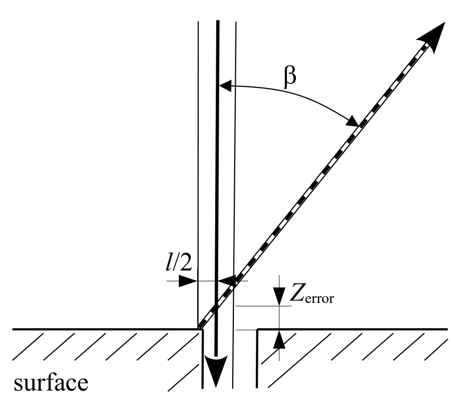
$$\begin{aligned} \tan\frac{\beta}{2} &= \frac{t}{2d} \\ d &= \frac{t}{2\tan\frac{\beta}{2}} \\ &= 4.94\,\text{cm} \end{aligned}$$

Credit was also given for answers presuming the laser was directed down the side of the hole, which gives a slightly lower maximum depth of $d = \frac{t}{\tan\beta} = 4.88\,\text{cm}$.

(b) (i) The surface is scanned optimally, which means that it should be as close as possible to the laser scanner. Hence the distance will be $Z = 26.6\,\text{cm}$. At this distance, the angle between the laser and reflection is given by:

$$\begin{aligned} \tan\beta &= \frac{L}{Z} \\ \beta &= 48.44^\circ \end{aligned}$$

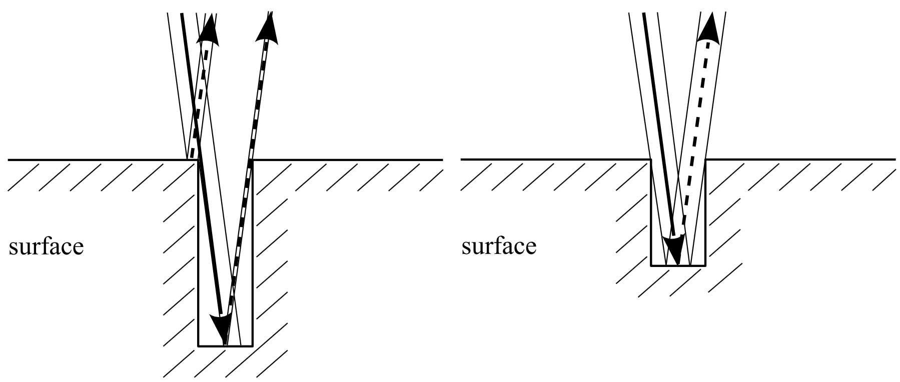
If the width of the laser is $l = 2\,\text{mm}$ and the maximum error is given by Z_{error}, then the setup is as below.



Hence

$$\tan\beta = \frac{l}{2Z_{\text{error}}}$$
$$Z_{\text{error}} = 0.9\,\text{mm}$$

(ii) The diagrams below show what happens when scanning the bottom of the hole with the laser.



If the hole is the same depth as before, the left-hand diagram shows that most of the reflection comes from the surface and we would not measure the hole in this situation. The right-hand diagram shows that all the laser has to reach and be reflected from the base of the hole to measure it correctly: effectively the new hole diameter can be considered to be $2\,\text{mm}$ less, and hence the maximum hole depth would be $4.94 \times (1 - 0.2)$ cm (numerical answer not required). We could detect a depth slightly greater than this so long as a large enough proportion of the laser reaches the base of the hole, but there would be an error (as in (i)) in that measurement.

(c) Other sources of error include:

Camera sensor size The depth resolution is determined by the pixel size in the camera image — more pixels give better resolution. The resolution also reduces with distance from the camera and laser, so scanning surfaces as close as possible reduces this error.

Surface properties Laser scanning relies on a single, clear, reflection of the laser off the surface. Surfaces which are highly specular, or translucent, or have fine features (e.g. hairs) will give multiple or distorted laser reflections. It is sometimes possible to make the surface more diffuse and simpler, for instance by wearing tight fitting clothes or dousing the surface in white powder.

Object movement It generally takes several minutes to scan most objects and the object must remain stationary during that time. Obviously, the faster the scan, the lower this error is likely to be. For a non-stationary object, it is better to rotate the laser scanner rather than the object.

Assessor's remarks: This was a fairly popular question on laser scanning. The first two parts in (a), worth nine marks in total, were standard calculations as given in lectures: many students got perfect marks for these sections, but there were too many who did not seem to know how to approach this at all. (b)(i) was the hardest part of the question and only a few got full marks for this, with many calculating errors due to camera resolution rather than due to laser thickness. (b)(ii) was answered better, with some good sketches and observations, including that the hole diameter was effectively reduced by the laser thickness. Most candidates came up with some other sources of error in (c), though too many included the thickness of laser light, which was not an "other" source of error.

4. **Texture mapping**

(a) 2D texture mapping involves associating pixel colours with a predefined image according to $(c_r, c_g, c_b) = \mathbf{T}(s,t)$. The programmer associates each polygon vertex (x_l, y_l, z_l) with a point (s,t) in the texture image: that vertex is then rendered with colour $\mathbf{T}(s,t)$. The colours of points inside the polygon are found by interpolating the texture coordinates (s,t) between the vertices, then indexing into the texture image $\mathbf{T}$. Perspective interpolation is required for perspective projections. Correct perspective 2D texture mapping is widely supported on commodity graphics hardware. It can be used for a variety of effects, including imparting realistic skin textures (freckles, hairs, scars etc) on computer-generated characters and reflecting the distant environment in shiny objects (*environment mapping*).

3D texture mapping is also supported in hardware but typically requires considerably more texture memory. The texture is now a 3D volume $(c_r, c_g, c_b) = \mathbf{T}(s,t,r)$ and the programmer associates vertex coordinates (x_l, y_l, z_l) with 3D texture coordinates (s,t,r). Inside the polygons, 3D perspective interpolation is used to generate the (s,t,r) values. One application is to accelerate volume rendering. [25%]

(b) This is clearly a perspective projection since $w = -z_v$. So we cannot interpolate linearly in screen space and must instead use perspective-correct interpolation. The easiest way to do this on paper is to transform back to view coordinates and establish how far P is from A and B in 3D view space. The homogeneous equations are relatively straightforward to solve if we set w to 1 on the left hand side and instead let the scale on the right hand side vary. Starting with vertex A:

$$\begin{bmatrix} 0 \\ 0.5 \\ 0.55 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1.1 & -0.11 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} kx_v \\ ky_v \\ kz_v \\ k \end{bmatrix}$$

$$\Leftrightarrow x_v = 0\ ,\quad ky_v = 0.5\ ,\quad -1.1kz_v - 0.11k = 0.55\ ,\quad -kz_v = 1$$

$$\Leftrightarrow 1.1 - 0.11k = 0.55 \Leftrightarrow k = 5\ ,\quad y_v = 0.1\ ,\quad z_v = -0.2$$

So in view coordinates A is $(0, 0.1, -0.2)$. Likewise for vertex B:

$$\begin{bmatrix} 0 \\ 0 \\ 0.99 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1.1 & -0.11 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} kx_v \\ ky_v \\ kz_v \\ k \end{bmatrix}$$

$$\Leftrightarrow x_v = 0 \ , \quad y_v = 0 \ , \quad -1.1kz_v - 0.11k = 0.99 \ , \quad -kz_v = 1$$

$$\Leftrightarrow 1.1 - 0.11k = 0.99 \Leftrightarrow k = 1 \ , \quad z_v = -1$$

So in view coordinates B is $(0, 0, -1)$. Finally for the rendered midpoint P:

$$\begin{bmatrix} 0 \\ 0.25 \\ 0.77 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1.1 & -0.11 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} kx_v \\ ky_v \\ kz_v \\ k \end{bmatrix}$$

$$\Leftrightarrow x_v = 0 \ , \quad ky_v = 0.25 \ , \quad -1.1kz_v - 0.11k = 0.77 \ , \quad -kz_v = 1$$

$$\Leftrightarrow 1.1 - 0.11k = 0.77 \Leftrightarrow k = 3 \ , \quad y_v = 1/12 \ , \quad z_v = -1/3$$

So in view coordinates P is $(0, 1/12, -1/3)$, which is $1/6$ of the way along the edge from A to B. The correct texture coordinates for P are therefore $1/6$ of the way from $(0, 0)$ to $(600, 0)$, which is $(100, 0)$. [40%]

(c) This is clearly an orthographic projection since $w = 1$. Since there is no perspective foreshortening, we simply need to linearly interpolate the texture coordinates in screen space. The texture coordinates of vertex P should therefore be $(300, 0)$. [10%]

(d) The renderings show a Phong-shaded surface mesh (red) alongside a CT image which is likely rendered as a 2D texture map on a single polygon. The frame rate will depend on both vertex (geometrical transformations) and pixel (shading and texture interpolation) operations, with the relative costs depending on the particular hardware and software implementation. The number of vertex operations is roughly the same in (a) and (b), since the mesh is not clipped, but there are more pixels to shade and interpolate in (b) which explains the lower frame rate. In (c), even though some of the mesh is clipped, expensive Phong-shading is required for roughly half the pixels in the viewport, so the frame rate drops again. In (d), the number of pixels to shade is around the same as in (c), but there is a lot more clipping and hence far fewer vertex operations, so the frame rate increases a little.

Note that the texture in (d) is very blocky. At this level of magnification, many screen pixels will map to the same texture pixel and it looks as if nearest neighbour texture interpolation has been chosen. The rendering would be improved by selecting linear or higher order interpolation instead. [25%]

Assessor's remarks: This question tested the candidates' understanding of texture mapping and surface rendering. The book work in part (a) was well answered by most, though a significant minority of candidates did not know what texture mapping was at all, despite it being the subject of almost an entire flipped lecture, a worked example, an examples paper question and many past tripos questions. In (b), few candidates were able to convert from 3D screen to view coordinates, and even fewer could then deduce the correct texture coordinate. In (c), only a handful of candidates identified the orthographic projection and, hence, the trivial texture coordinate. In (d), many candidates recalled the frame rate effects from the lab, though there was again a sizeable minority who appeared to have not engaged with this aspect of the course at all.

Andrew Gee & Graham Treece
May 2025